From:	Eric Blischke
To:	rjw@nwnatural.com; johnt@windwardenv.com
Cc:	Chip Humphrey; Burt Shephard; rgensemer@parametrix.com; jay.field@noaa.gov; Joe Goulet
Subject:	
Date:	07/17/2009 05:39 PM

Bob, As you are aware, we have been discussing some of the details of the LWG's interpretation of the Portland Harbor sediment bioassay results. Some elements of the interpretation were discussed during a conference call on Thursday, June 18, 2009. EPA has further reviewed the bioassay results and information regarding the bioassay interpretation provided by the John Toll on behalf of the LWG.

Here is where I believe we are:

1) No transcription errors were identified during a review of the reference envelope bioassay results. However, minor discrepancies between effect levels observed between EPA and LWG have been noted in the raw data. EPA believes these are due to rounding errors and use of different numbers of significant digits during calculations. To eliminate these discrepancies, calculations based on the raw toxicity data, such as the proportion of sample survival and biomass relative to laboratory control and the reference envelope calculations, should be carried through to as many significant digits as the data permits. Rounding of results should take place only after calculations are complete
2) The total biomass calculations were done correctly. Although small discrepancies were identified, they appear to be due to either minor differences in rounding or significant digits or, in the case of the two stations with duplicate results, the treatment of these samples as individual samples rather than the arithmetic mean of two data points.
3) Mortality should be calculated as percent survivorship relative to control and control normalized values should be computed as test/control. This is consistent with Table 2-1 in the March 17, 2006 Bioassay Interpretation Report, ASTM Method E-1706, and EPA Guidance.
4) Duplicate reference envelope samples should be pooled (arithmetic mean) rather than treated as individual samples. This is consistent with EPA's March 31, 2009 direction on the development of the reference envelope (see last paragraph, first page).
5) Identification of Level 1, Level 2 and Level 3 thresholds: The toxicity thresholds should be calculated based on 10% of the reference envelope not an absolute 10%. This is consistent with Tables RE 1, RE-2 and the text of EPA's March 31, 2009 direction on the Calculation and Use of Reference Envelope for Portland Harbor Sediment Toxicity Test Interpretation.
6) The reference envelope for the four types of sediment toxicity tests performed at Portland Harbor is the lower 5th percentile of the best fitting statistical distribution of the survivorship and biomass responses of the test organisms at the 17 reference envelope sample locations. The 5th percentile will be calculated as follows:

- The raw toxicity data are to be expressed in terms of the proportion of survivorship or biomass, depending on the toxicity test, relative to the laboratory control survivorship or biomass. Reference envelope responses for each of the 17 stations are to be calculated using the following formulas: Survival = T / C, where T and C are the test sample (T) survival and laboratory control sample (C) survival, respectively. Biomass = REB / LCB, where REB and LCB are the reference envelope

sample (REB) and laboratory control sample (LCB) biomass, respectively.

- The survival and biomass results for all 17 reference envelope station for each of the four toxicity tests are placed into rank order.
- Using a statistical software package to be approved by EPA, the best fitting statistical distribution for each of the four sets of reference envelope data will be calculated. The specific software package to be employed is not as important as it being one of the more statistically robust packages with a number of distributions available for fitting. Examples of acceptable software packages include, but are not limited to SAS, Systat, SPSS, Statistica, Best Fit, @Risk or Crystal Ball.
- Several statistical procedures for quantifying the fit of a given distribution (e.g. normal, lognormal, logistic, Weibull, gamma, etc.) are available. Among the more commonly employed procedures are the Anderson-Darling procedure, the Kolmogorov-Smirnov D statistic and the chi-squared test. Of these, the Anderson-Darling procedure is preferred by EPA, because it gives more weight to the fit in the tails of the distribution than either the Kolmogorov-Smirnov statistic or the chi-squared test.
- It is likely that more than one statistical distribution will give high quality fits to each of the four reference envelope data sets. As a check on the statistical distribution fitting output, probability plots for the three best fitting distributions for each of the four reference envelope data sets will be plotted. With EPA agreement, a distribution may be selected for a reference envelope data set that does not have the best fit based on the results of the Anderson-Darling procedure, but which, upon visual inspection of the probability plots, fits the lower tail of the reference envelope distribution better than the distribution with the best Anderson-Darling fit. It is reiterated that EPA must agree to the selected distribution for each of the four reference envelope data sets.
- Once the statistical distributions for the four reference envelope data sets are selected, the lower 5th percentile of each data set is calculated, using output from the same statistical package used to fit the distribution.

The above procedures for computing the results of the bioassay tests, calculating hit/no-hit designations, developing the reference envelope and identifying Level 1, Level 2 and Level 3 toxicity hits should be followed.

As we have discussed, it would be helpful if we can schedule a conference call with you and John to discuss this further.

Thanks, Eric